# Forecasting with quantitative methods the impact of special events in time series

Nikolopoulos, Konstantinos


**Forecasting with quantitative methods the impact of special events in time series**

| | |
|---|---|
| Journal: | *Applied Economics* |
| Manuscript ID: | APE-06-0447.R1 |
| Journal Selection: | Applied Economics |
| Date Submitted by the Author: | 31-Jul-2007 |
| Complete List of Authors: | Nikolopoulos, Konstantinos |
| JEL Code: | C10 - General < C1 - Econometric and Statistical Methods: General < C - Mathematical and Quantitative Methods, C15 - Statistical Simulation Methods\|Monte Carlo Methods < C1 - Econometric and Statistical Methods: General < C - Mathematical and Quantitative Methods, C45 - Neural Networks and Related Topics < C4 - Econometric and Statistical Methods: Special Topics < C - Mathematical and Quantitative Methods, C53 - Forecasting and Other Model Applications < C5 - Econometric Modeling < C - Mathematical and Quantitative Methods, D83 - Search\|Learning\|Information and Knowledge < D8 - Information and Uncertainty < D - Microeconomics |
| Keywords: | Special Events, Forecasting , Quantitative methods, Artificial intelligence, Expert methods |
| | |

Editorial Office, Dept of Economics, Warwick University, Coventry CV4 7AL, UK

# Forecasting with quantitative methods the impact of special events in time series


## Abstract

Quantitative methods are very successful in producing baseline forecasts of time series; however these models fail to forecast neither the timing nor the impact of special events such as promotions or strikes. In most of the cases the timing of such events is not known so they are usually referred as shocks (economics) or special events (forecasting). Sometimes the timing of such events is known a priori (i.e. a future promotion); but even then the impact of the forthcoming event is hard to estimate. Forecasters prefer to use their own judgment for adjusting for forthcoming special events, but humans' efficiency in such tasks has been found to be deficient. This study after examining the relative performance of Artificial Neural Networks, Multiple Linear Regression and Nearest Neighbor approaches proposes an expert method which combines the strengths of regression and artificial intelligence.

*Forecasting with quantitative methods the impact of special events in time series*

## 1. Introduction

There is a fundamental difference between a shock and an expected/forthcoming irregular event: a non-periodic/ event that could be announced and therefore expected at a given time (e.g. sales promotion activity). The term *shock* is used, predominantly in the economics literature, for events that are not expected at all (e.g. an earthquake, a stock-market crash) - thus the timing of a shock is not known a-priori. When it comes to shocks there are two issues to address: when the shock is going to happen and what is going to be the impact of it? As far as forthcoming irregular events are concerned, the only interest is on the potential impact of such an event on the baseline forecast. This latter category of events is usually mentioned in the forecasting literature as *special events* (Armstrong 2001).

The current study examines the relative performance of Artificial Neural Networks, Multiple Linear Regression and Nearest Neighbor approaches for estimating the impact on baseline series of expected irregular future events. Furthermore, building on the results of this comparison, an expert method is proposed which combines the strengths of regression and neural networks.

## 2. Literature review

In the forecasting literature, quantitative models have generally been very successful in extrapolating the basic trend-cycle component of business and economics time series (Assimakopoulos and Nikolopoulos 2000, Allen and Salim 2005, Halkos and Kevork 2006, Fildes *et al.* 1998, Jaesun and Kim 2006, Makridakis *et al.* 1982, Makridakis and Hibon 2000, Maris et. al. 2004&2007, Petropoulos *et. al.* 2005&2007, Sadorsky 2005). However, time series models have been found to be deficient when it comes to handling *special* events, which are non-periodic such as promotions, announcements, changes in regulations and strikes (Lee *et al.* 2007, Nikolopoulos *et. al.* 2007a)

*Forecasting with quantitative methods the impact of special events in time series*

Special events introduce an additive or multiplicative impact in the trend-cycle component (Webby R and O'Connor 1996). Both the literature and common practice suggest that forecasters tend to use their own judgment to adjust time series for irregular events (Armstrong 2001). However, the accuracy of these adjustments is often low (Goodwin 2000, 2002).

Forecasting with neural networks has been very popular in academic circles in the last decade (Aiken 1999, Heravi *et al.* 2004, Curry 2004, Kim 2003, Sanders and Manrodt 2003, Wang *et al.* 2003, Yao *et. al.* 2000, Zhang *et al.* 1998, Zhang 2001). While batch forecasting was not one of the success areas for neural networks (Balkin and Ord 2000, Makridakis and Hibon 2000), the estimation of the impact of irregular events proved to be a very fertile area of research. In particular, Lee and Yum (1998) have proposed a formal framework for judgmental adjustments in time series using Artificial Neural Networks (ANN) in order to estimate the additive impact of irregular events.

This framework has been enhanced by Nikolopoulos and Assimakopoulos (2003), to incorporate a variety of supportive models, such as Multiple Linear Regression (MLR – Makridakis *et al.* 1998) or Nearest Neighbors approaches (Härdle 1992), as well as to be able to adjust to take into account any kind of business-related irregular events. However this feature was not properly evaluated in due to the lack of real special events data.

More recently Nikolopoulos *et al.* (2006b) have examined a case study with TV ratings data where the use of ANNs proved quite competitive to unaided judgment, MLR and Nearest Neighbor approaches. Furthermore in a follow up study simulation results have been submitted indicating that neither linear nor non-linear methods can effectively used solely, as both has advantages and disadvantages when applied to certain types of special events (Nikolopoulos *et. al.* 2007a, b)

## 3. A framework for adjusting for past and future special events

A successful forecasting function within an organization is usually conditional on the participation of a variety of *experts*. A *forecasting expert* should be able to spot the appearance of an irregular event in the past of a series; this person will be the *forecasting manager*. In addition, *marketing experts and production managers* should be able to identify the event as well as the parameters that affected it and potentially quantify them. For example a forecasting expert should spot that the sales in a specific period in the past are more than expected for that period of the year, while the marketing expert should be able to identify the promotional activity that took place in this period and be able to recall the dimensions of that event (e.g. budget spend, media used, duration).

Ideally, the impact of such an event in the past would be quantified and removed from the series resulting in a filtered series that was more appropriate for extrapolation (Armstrong 2001). On the other hand, the future appearance of irregular events could be known in advance, and if an expert could forecast or even know the values of the event's parameters, an appropriate model could estimate the potential impact on the baseline series. A baseline series is defined as a filtered series embedding only trend and cycle components (Makridakis *et al.* 1998).

Lee and Yum (1998) have proposed all the necessary entities so as to describe, adjust and forecast the impact of an irregular event. An expert in the specific market (someone possessing all the necessary domain knowledge for the time series under consideration) should:

- <u>Define Judgmental Factors</u>; the expert should be able to identify *promotions* as a judgmental factor - that is an irregular factor that can influence the time series of historic sales for a specific product.

- Identify Historical Judgmental Events; the expert should be able to identify when exactly an irregular event has happened as well as to define which factors characterize the event (and the values of the parameters for each factor). For example, a promotion could have been applied from May 2001 for two months spending in total 10.000€ in TV advertisements.

- Break events' impact into periods; an expert should be able to split the events into instances per period (week, month, etc) and define the impact on the baseline series for each instance either additively or multiplicatively. For example, a promotion might have been applied from May 2001 for two months with an impact of a 12% increase in the sales of May and an increase of 3% in June, respectively (relatively to the average sales of those months in the previous years).

- Remove the Impact of Judgmental Instances from Time-series; an expert or an automated FSS should be able to remove the impact of the judgmental instances of the irregular events from the historic data. As a result a filtered series should be created, a series that is far more appropriate for extrapolation via time series models.

On the expectation of a future irregular event, an expert should adjust his statistical forecasts (Armstrong 2001). This is a very tricky procedure that potentially introduces major errors into the forecasting process due to biases and inefficiencies of the adjusters (Goodwin 2000, 2002). Thus, there is need for automated support from FSS so as to enhance this adjustment process (Nikolopoulos and Assimakopoulos 2003). The forecaster/expert should be able to:

- Record expected irregular events; for example, the expert should know when the next promotion is due to take place, as well as all the relevant information that relates to the promotion.

- Build a model for judgmental adjustment; a model should be built so as to estimate the impact of the future events, based on the historic impacts of similar events in the past. There are very few studies on this topic. Those that have been carried out mainly refer to case studies where a very specific time-series is examined (e.g.. oil sales in country X). Methods used from time to time include complex methods such as ANN (Lee and Yum 1998) as well as approaches such as the Delphi method (Rowe G, Wright 1999) and AHP (Flores et al. 1992).

- Apply adjustments to the  baseline extrapolation; forecasts produced by the model should be superimposed with the baseline forecasts so as to constitute the final forecasts.

## 4. Forecasting the Impact of future irregular events

In this section we discuss four different approaches that seem appropriate for forecasting the impact of irregular events. The events are characterized by a single variable standing for the impact of the event accompanied by a set of cues that an act as independent variables.

### 4.1 Nearest Neighbors

Nearest Neighbors (NN) approaches (Härdle 1992), are a very simple way of estimating the impact of events. They involve looking at the past in order to assess the impact of similar events (Green 2002).

*Finding the neighbors*

This task requires ranking historic events in terms of similarity. Similarity requires a metric of distance. Various metrics have been used in order to measure distance in the $n$D-space as Euclidian norms (Härdle 1992). $n$D-space metrics are required since our events involve in general more than one attribute. In this study a very straightforward metric, based on  the

Absolute Percentage Error (APE - Makridakis *et al.* 1998), was used to measure the distance between similar events,

$$\left\| Ev1 - Ev2 \right\| = \sum_{i=1}^{n} \left| \frac{P_{i,Ev1} - P_{i,Ev2}}{P_{i,Ev2}} \right|$$, where $Ev$: event, $P$: attribute value.

For example, in the case of a promotion characterized only by *budget* and *duration,* the above formula becomes:

$$\left\| Ev1 - Ev2 \right\| = \sum_{i=1}^{2} \left| \frac{P_{i,Ev1} - P_{i,Ev2}}{P_{i,Ev2}} \right| = \left| \frac{Budget_{Ev1} - Budget_{Ev2}}{Budget_{Ev2}} \right| + \left| \frac{Duration_{Ev1} - Duration_{Ev2}}{Duration_{Ev2}} \right|$$

The lower the value of this measure, the more similar the events.

*Estimating Impact*

Two different methods will be used in order to estimate the impact of future events:

- BN

  This is the simplest approach, where the impact of the nearest neighbor provides the impact for the future event.

- B3N

  Here, the three nearest neighbors are used for the estimation of the impact of the future event. The final impact is constructed as 50% of the impact of the nearest neighbor and 25% from the second and third nearest neighbors, respectively. In mathematically terms, a triangular kernel function is used (Härdle 1992), assigning weights equal to (1/2, 1/4, 1/4) to the three nearest neighbors.

### 4.2 Regression

Classical Multiple Linear Regression (MLR – Makridakis *et al.* 1998) can be used in order to estimate the impact of the holdout events:

$$impact = c + \sum_{i=1}^{n} b_i P_i \left| \frac{P_{i,Ev1} - P_{i,Ev2}}{P_{i,Ev2}} \right|$$

For example, in the case of a promotion characterized only by *budget* and *duration,* the above formula becomes:

$$impact = constant + b_1 * \text{Budget} + b_2 * \text{Duration}$$

The model parameters are estimated from the values of the cues for past events using the least squares criterion.

### 4.3 Artificial Neural Networks

One of the simplest topologies of ANN, a fully connected multilayer perceptron with two hidden layers and an output layer can be used for the estimation of the impact of irregular events (Haykin 1998); the input signals would be the event's attribute values while the output would be the impact of the event under consideration (Figure 1).

**[Insert Figure 1 about here]**

In this study the size of the input signal is equal to the maximum number of parameters explaining the historic irregular events. In order to forecast the impact of a possible future event, the network is trained first with all historic data as the training set, using the Back Propagation (BP) algorithm (Haykin 1998). Each historic event is considered to be one training example while the impact is the response of the network. As soon as the training of the network is completed, and the weights of the nodes have been determined, the future event's parameters are fed into the network with the form of an input signal. For each parameter in the input signal that does not appear in the future event, a zero value is assigned. Signal flow through the network progresses in a forward direction, from left to right and on a layer-by-layer basis reaching the output node. The outcome is the potential impact of the future event.

*Training*

All different fully connected networks up to two hidden layers and no more neurons per hidden layer than the input size are trained ten times each. The number of alternative networks is limited by the constraint that the second hidden layer is not used if the first hidden layer does not contain neurons equal in number with the input size. Each time, the Mean Square Error (MSE) over the training sample is calculated (Makridakis *et al.* 1998). The network with the smallest MSE error is used to produce the forecast.

*ANNs and MLR*

There are similarities between ANNs and MLR. Neural networks can be seen as a form of nonlinear regression (Ripley 1994). It has also been observed that MLR can be expressed as a simple ANN with only one layer (Warner and Misra 1996). However, even if MLR and an ANN model have exactly the same topology, the weights that would be given would only be identical by chance as there are thousands of training methods suggested in the ANN literature.

### 3.4 Expert method

Our proposal for an expert approach is based on the assumption that linear as well as non-linear relationships exist between an event's impact and the values of its attributes. MLR is expected to be better at capturing the linear relationships, while ANN should be superior at capturing the non-linearities, as their topology permits the adaptation to more complex relationships and interactions. We therefore propose a selective algorithm which chooses between MLR and ANN according to whether nonlinearities are detected.

How can we detect whether the relationship: Impact = *f*(attribute values) is linear or not? We propose two different approaches:

- EXPRT1

  Here the MLR model is fitted to the past events and $R^2$ is calculated. If $R^2 > 95\%$, the relationship between the impact and its determinants is considered linear and MLR is used to estimate the impact of the future event. Otherwise ANN is used for that estimation.

- EXPRT2

  Since the noise of the series is not known *a priori*, it might be inadvisable to set up a threshold as high as 95% for $R^2$. In this approach the MLR model is fitted to the past events and $R^2$ is calculated. The ANN is trained with the past events and $R^2$ is also calculated (Since estimations for all past events can be calculated with the optimized synaptic weights, $R^2$ can be calculated). If $R^2_{MLR} > R^2_{ANN}$ the relationship between the impact and its determinants is considered linear and MLR is used to estimate the impact of the future event. Otherwise ANN is used for that estimation.

## 4. Evaluation

This section covers the results of the simulation experiment used to evaluate the alternative forecasting methods as well as evidence from an application to real life data via a case study in forecasting TV ratings.

### 4.1 Simulation setup

*The Data Sets*

Each data set is constructed from 50 special events – the 35 first are used as a learning set and the remaining 15 used for testing out of sample accuracy - measured in terms of the Mean Absolute Percentage Error (MAPE) and the Median Absolute Percentage Error (MdAPE) (Makridakis *et. al.* 1998). So the first 35 special event can be seen as past events while the remaining 15 as forthcoming/announced special events withi known parameters however unknown impact.

The 50 special event correspond to non-consecutive points of a time series i.e. you could have a monthly time series with 10 years worth of data (120 months) where only 50 of those months having promotions (the most common type of special event in operational forecasting); 35 in the first 7 years (the timings randomly chosen) and 15 in the last three years that were held out for evaluation. Thus you have 70 periods with non-promotial activity and 50 with (35 and 15 within and out-of-sample respectively). It is out of the scope of this study to measure the accuracy in periods where special events are taken place, We do take a decomposition stance and try to estimate/isolate the additive impact of the special event in the baseline level of the series. So a partial forecast is the target, which is exactly the isolated impact of the special event (figure 2).

**[Insert Figure 2 about here]**

*Factors-Parameters*

Two factors are considered. The first one, <u>Promotion</u>, has a positive effect and is described via three attributes:

a) <u>Budget</u>, ranging from 50 to 150 in steps of 10 (each step representing 1000€),

b) <u>Duration</u>, ranging from 1 to 14 days,

c) <u>Media</u> used, with values: 1 (Paper), 2 (Paper + Radio), 3 (Paper + Radio + TV)

The second factor, Strike, has a negative effect and is described via two attributes:

a) Percentage, ranging from 20% to 100% in steps of 5% (representing the % participation of employees in the strike),

b) Duration, ranging from 1 to7 days.

The rational behind the building of the special events comes from the marketing literature (DelVecchio *et al.* 2006, Martvnez-Ruiz *et al.* 2006, Laroche *et al.* 2003, Kotler 2003); where typically a positive event has a positive bell-shaped impact (figure 3a), fading after a certain number o periods (vice versa for a negative one). Sometimes the promotion creates a skewed bell- shaped impact when if forces sales faster (figure 3b) or slower; furthermore sometimes after some periods there is a negative effect in the sales as the product has been oversold during the promotional activity (figure 3c). For the sake of simplicity and to reduce the degrees of freedom of this multiple hypothesis setup, I only investigate special events with impact in one period only – where that impact has a pulse shape so in an additive model it is just a straight addition (or deduction) to the regular sales (regular standing for an equivalent period without promotional activity i.e. last year's sales for the same period in the year)

**[Insert Figure 3 about here]**

*Linearity*

For the construction of the test events we used the following generating function:

*Impact = f(Factor attribute values) + Noise*

For the function *f,* two formulae were used: Linear and Non-linear as shown below:

Promotion/Linear : *0.7 x Budget + 5 x Duration + 20 x Media*

Strike/Linear : - *100 x Percentage – 2 x Duration*

Promotion/Non Linear : *(0.5 x Budget) x ((Duration/14) + Media)*

Strike/Non Linear : - *50 x Percentage x ((Duration/7)+1)*

*Noise*

Two levels of Gaussian noise $N(\mu,\sigma^2)$ were generated: Low N(0,10) and High N(0,30)

*Runs*

We ran 50 simulations for each of the four combinations of *Linearity* and *Noise*, resulting in 200 simulations.

### 4.2 Simulation Results

The 200 datasets were analyzed, each one consisting of 50 events. 35 events were used to estimate the models' weights and 15 were used for testing. Thus the overall accuracy calculation was based on a sample of 3000 events. The results are presented in tables 1 to 4.

**[Insert Table 1 about here]**

Overall, EXPRT1 was the better approach, though its performance was not much better than using ANN for all events. Overall ANN was the best among the first four standard approaches and it was significantly better than MLR. This is indicative of the strong effect of the simulated nonlinearities. Both expert approaches were better than all the other methods. The use of the MdAPE, rather than the MAPE, did not alter the ranking of the methods.

**[Insert Table 2 about here]**

In table 2 the effect of linearity is presented. As expected, for the linear subset, MLR performed best, while for the non-linear one ANN was the top performer. EXPRT1 and EXPRT2 were close to the best methods in each case. Overall, EXPRT1 was the better approach. However its performance did not differ much from ANN. As before, the use of the MdAPE did not alter these results.

**[Insert Table 3 about here]**

In table 3 the effect of noise is examined. For all levels of noise the EXPERT1 and EXPRT2 approaches were better than the conventional methods.(Again, the use of the MdAPE made little difference).

**[Insert Table 4 about here]**

Table 4 provides the computation times for each model, a criterion critical nowadays where retailers have to produce forecasts for thousand of products every day. As expected, ANN and, inevitably, the expert approaches required significantly more processing time.

### 4.3 An application to real-life data

Data from a real-life case study are used to test the aforementioned models and refer to television (TV) audience ratings from 1996 to 2000 in Greece (Nikolopoulos *et. al.* 2007b). TV audience ratings are measured as a percentage (%) of people watching a specific programme; when the TV programme broadcasts a special event, such as a major sport event the audience level is radically increased. Impact refers to the additive increase in the percentage of the total audience above the average of the last four relative periods (same time zone within day, same day of week) for a given channel.

AGB Hellas S.A. provided for this study real TV audience ratings data, relating to the target group of men aged 25-44 in Greece. The supplied data included 46 sport programmes. Each event is described by three categorical parameters: Importance, Competition and Time Zone. The parameters are quantified on scales from 1 to 3 or 1 to 5 in order to make statistical forecasting models easily applicable.

From the 46 programmes for which data was provided, 34 were used for the fitting or training the models and 12 were used as a hold-out sample to evaluate forecasting performance. The values for the three independent variables have been estimated by an expert in audience analysis in Greece. A subset of this data is given in table 5.

**[Insert Table 5 about here]**

The performance of MLR, BN, B3N, ANN was tested on these real life data and accuracy was measured in terms of Mean Absolute Percentage Error. There was no need for a Percentage/scaled error in this case study as all cases was within the same scale; thus the simplest of unsigned error metrics was chosen;

**[Insert Table 6 about here]**

In this real-life example ANN performs better indicating strong non-linearities in the data. Nearest Neighbor Approaches follow while MLR comes last; no expert approaches were separately reported in this table as the selection protocols EXPRT1 and EXPRT2 were selecting in all 12 cases the ANN as to perform extrapolation, thus the results were identical to the ANN one (in this real-life example the ANN is trained only once in the available past 34 cases).

## 5. Conclusions and Future Research

This study has examined the relative performance of ANN, MLR and NN in estimating the impact of expected irregular future events. The simulation results suggest that the use of two expert approaches should be considered, based on selection protocols between ANN and MLR.

There is strong evidence that, overall, the expert approaches do perform significantly better. However, the simpler the problem (linear, low noise) the greater the advantage for regression approaches. For more complex problems ANN has a critical advantage. Thus, it is no surprise that, overall, a selection protocol outperformed all other approaches as long as nonlinearities could be successfully detected. For this detection two different approaches were tested, EXPRT1 and EXPRT2. Although EXPRT1 performed slightly better than EXPRT2, we do consider EXPRT2 to be a more promising approach as it is not dependent on the preset threshold of 95% that might have given EXPRT 1 an advantage with the specific dataset that was examined in this study.

Further results from an application of these approaches to real-life data of forecasting TV ratings of major sport events showed promising evidence but at the same highlight strong non-linearities and thus revealing how much more difficult is to forecast real-life data (contrasting to simulating data where always an underlying stochastic mechanism creates the datasets).

Future research should focus on constructing an expert model that both detects nonlinearities more efficiently and also makes it selection from a larger set of forecasting methods. Also focus should be given as to how to reduce the computational intensity and thus the required time to setup, optimize, train and run the Artificial Neural Network approaches as to make the applicable in real-life setup. Lastly some more effort should be given on the improvement of the performance of Nearest Neighbor approaches as there are quite intuitively appealing.

Concluding, it should be noted that the severe impact and frequency of irregular events in practical forecasting contexts means that there is a widespread need for models to support judgmental forecasters in their task of adjusting baseline forecasts to take into account the effects of these events. The search for effective modeling procedures is therefore likely to be a fertile research field for many years to come and the synergy required from the Information Technology is rather crucial (via Data Mining and Information Management techniques and algorithms – Fildes and Nikolopoulos 2006).

**Acknowledgements**

The Manchester Business School (MBS) Research Support fund 2006/7 funded the work for this paper as part of a one year interdisciplinary project on the development of an innovative forecasting methodology for television ratings. The author would like to thank especially Mr. Konstantinos Giannelos for providing the software for running the experiments with the ANNs, a software developed in July 2004 in the Forecasting Systems Unit in National Technical University of Athens for his Diploma dissertation entitled “Irregular Events Impact Estimation Methodology and Information System”, under the supervision of Dr K Nikolopoulos, Assistant Professor in MBS and Professor Vassilis Assimakopoulos, Special Secretary for Digital Planning in the Ministry of Economics and Finance in Greece. The author would like to thank Professor Paul Goodwin of the School of Management at the University of Bath for his suggestions in the first submission of this research as well as the Editor of AE and an anonymous referee for their suggestions in the revision of this article.

## References

Aiken, M. (1999), "Using a neural network to forecast inflation", *Industrial Management and Data Systems*, Vol. 99 No. 7, pp. 296-301.

Allen, D.E. and Salim, H.M. (2005), "Forecasting profitability and earnings: a study of the UK market (1982-2000)", *Applied Economics*, Vol. 37 No. 17, pp. 2009–2018.

Armstrong, J.S. (2001), *Principles of Forecasting: A handbook for researchers and practitioners*, Kluwer Academic Publishers, Dordrecht.

Assimakopoulos, V. and Nikolopoulos, K. (2000), "The theta model: a decomposition approach to forecasting", *International Journal of Forecasting*, Vol. 16 No. 4, pp. 521-530.

Balkin. S. and Ord, K. (2000), "Automatic Neural Net Modelling for univariate time series", *International Journal of Forecasting*, Vol. 16 No. 4, pp. 509-515.

Curry, B. (2004), "'Simple' neural networks for forecasting", *Omega* Vol. 32 No. 2, pp. 97-100.

DelVecchio, D., Henard, D. H. and Freling, T. H. (2006), "The effect of sales promotion on post-promotion brand preference: A meta-analysis", *Journal of Retailing*, In Press, Corrected Proof, Available online 30 June 2006,

Fildes, R., Hibon, M., Makridakis, S. and Meade, N. (1998), "Generalising about univariate forecasting methods: further empirical evidence", *International Journal of Forecasting*, Vol. 14 No. 3, pp. 339-358.

Fildes, R. and Nikolopoulos, K. (2006). "Spyros Makridakis: An interview with the International Journal of Forecasting", *International Journal of Forecasting*, Vol. 22 No. 3, pp. 625-636.

Flores, B.E., Olson, D.L. and Wolfe, C. (1992), "Judgmental adjustment of forecasts: A comparison of methods", *International Journal of Forecasting*, Vol. 7 No. 4, pp. 421-433.

Goodwin, P. (2000), "Improving the voluntary integration of statistical forecasts and judgment", *International Journal of Forecasting*, Vol. 16 No. 1, pp. 85-99.

Goodwin, P. (2002), "Integrating management judgment and statistical methods to improve short-term forecasts", *Omega,* Vol. 30 No. 2, pp. 127-135.

Green, K.C. (2002), "Forecasting decisions in conflict situations: a comparison of game theory, role-playing, and unaided judgement", *International Journal of Forecasting,* Vol. 18 No. 3, pp. 321-344.

Härdle, W. (1992), *Applied Nonparametric Regression* (Econometric Society Monographs), Cambridge University Press, Cambridge.

Halkos, G.E. and Kevork, I.S. (2006), "Forecasting the stationary AR(1) with an almost unit root", *Applied Economics Letters*, Vol. 13 No. 12, pp. 789–793.

Haykin, S. (1998), *Neural Networks: A Comprehensive Foundation* (International Edition). Pearson US Imports & PHIPEs.

Heravi, S., Osborn, D.R. and Birchenhall, C.R. (2004), "Linear versus neural network forecasts for European industrial production series", *International Journal of Forecasting*, Vol. 20 No. 3, pp. 435-446.

Kim, K-J. (2003), "Financial time series forecasting using support vector machines". Neurocomputing, Vol. 55 No. (1-2), pp. 307-319.

Kotler, P. (2003), *Marketing Management* 11th Edition, Prentice-Hall, New York.

Laroche, M., Pons, F., Zgolli, N., Cervellon, M.-C. and Kim, C. (2003), "A model of consumer response to two retail sales promotion techniques", *Journal of Business Research*, Vol. 56 No. 7, pp. 513-522.

Lee, J.K. and Yum, C.S. (1998), "Judgmental adjustment in time series forecasting using neural networks", *Decision Support Systems*, Vol. 22, pp. 135-154.

Lee, W.Y., Goodwin, P., Fildes, R., Nikolopoulos, K. and Lawrence, M. (2007), "Providing support for the use of analogies in demand forecasting tasks", *International Journal of Forecasting*, Vol. 23 No. 3 (forthcoming)

Makridakis, S., Andersen, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., Newton, J., Parzen, E. and Winkler, R. (1984), *The forecasting accuracy of major time series methods*, John Wiley & Sons, New York.

Makridakis, S., Wheelwright, S., and Hyndman, R. (1998), *Forecasting, Methods and Applications*, 3rd Edition, Wiley, New York .

Makridakis, S. and Hibon, M.  (2000), "The M3-Competition: Results, conclusions and implications", *International Journal of Forecasting*, Vol 16 No. 4, pp. 451-476.

Maris, K., Pantou, G., Nikolopoulos, K., Pagourtzi, E. and Assimakopoulos, V. (2004), "A study of financial volatility forecasting techniques in the FTSE/ASE 20 index", *Applied Economics Letters*, Vol. 11 No 7, pp. 453-457.

Maris, K., Nikolopoulos, K., Giannelos, K. and Assimakopoulos V. (2007) "Options trading driven by volatility directional accuracy", *Applied Economics*, Vol. 39 No 2, pp. 253-260.

Martvnez-Ruiz, M.P., Mollα-Descals, A. Gσmez-Borja, M.A. and Rojo-Alvarez, J.L. (2006), "Using daily store-level data to understand price promotion effects in a semiparametric regression model", *Journal of Retailing and Consumer Services*, Vol. 13 No. 3, pp. 193-204.

Nikolopoulos, K. and Assimakopoulos, V. (2003), "Theta Intelligent Forecasting Information System", *Industrial Management and Data Systems,* Vol. 103 No. 9, pp. 711-726.

Nikolopoulos, K., Bougioukos, N., Giannelos, K. and Assimakopoulos V. (2007a) "Estimating the impact of Shocks with Artificial Neural Networks", *Lecture Notes in Computer Science*, Vol. 4669, pp. 516-526.

Nikolopoulos, K., Goodwin, P., Patelis, A. and Assimakopoulos, V. (2007b) "Forecasting with cue information: a comparison of multiple regression with alternative forecasting approaches". *European Journal of Operations Research ,* Vol. 180 No. 1, pp. 354-368.

Noh, J. and Kim, T-H (2006), "Forecasting volatility of futures market: the S&P 500 and FTSE 100 futures using high frequency returns and implied volatility", *Applied Economics*, Vol. 38 No. 4, pp. 395–413.

Petropoulos, C., Nikolopoulos, K. Patelis, A., Assimakopoulos, V. and Askounis, D. (2007), "Tourism Technical Analysis System", *Tourism Economics*, Vol. 12 No. 4, pp. 543-563.

Petropoulos, C., Patelis, A., Nikolopoulos, K. and Assimakopoulos, V. (2005), "A technical analysis approach to tourism demand forecasting", *Applied Economics Letters*, Vol. 12 No. 6, pp. 327-333.

Ripley, B.D. (1994), "Neural Networks and Related Methods for Classification", *Journal of the Royal Statistical Society*, Series B, Methodological Vol. 56, pp. 409-456.

Rowe, G. and Wright, G. (1999), "The Delphi technique as a forecasting tool: issues and analysis", *International Journal of Forecasting*, Vol. 15 No. 4, pp. 353-375.

Sanders, N.R. and Manrodt, K.B. (2003), "The efficacy of using judgmental versus quantitative forecasting methods in practice", *Omega,* Vol. 31 No. 6, pp. 511-522.

Sadorsky, P. (2005), "Stochastic volatility forecasting and risk management", *Applied Financial Economics*, Vol. 15 No. 2, pp. 121–135.

Wang, W., Lu, W., Wang, X. and Leung, A.Y.T. (2003), "Prediction of maximum daily ozone level using combined neural network and statistical characteristics", *Environment International,* Vol. 29 No. 5, pp. 555-562.

Warner, B., Misra, M. (1996), "Understanding Neural Networks as Statistical Tools", *The American Statistician*, Vol. 50, pp. 284-293.

Webby, R. and O'Connor, M. (1996) "Judgemental and statistical time series forecasting: a review of the literature", *International Journal of Forecasting*, Vol. 12 No. 1, pp. 91-118.

Yao, J., Li, Y. and Tan, C.L. (2000), "Option price forecasting using neural networks", *Omega,* Vol. 28 No. 4, pp. 455-466.

Zhang, G.P., Patuwo, B.E. and Hu, M.Y. (1998), "Forecasting with artificial neural networks: The state of the art", *International Journal of Forecasting*, Vol. 14 No. 1, pp. 35-62.

Zhang, G.P. (2001), "An investigation of neural networks for linear time-series forecasting", *Computers & Operations Research*, Vol. 28 No. 12, pp. 1183-1202.

# TABLES & FIGURES

**Figure 1. M**ultilayer perceptron (two hidden layers)

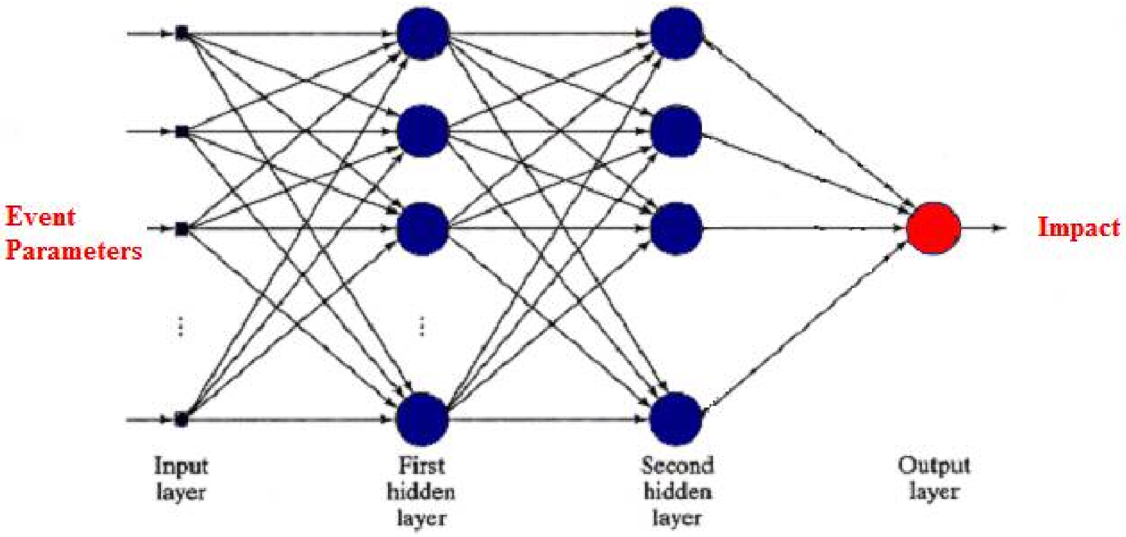

**Figure 2.** Promotional periods in a time series

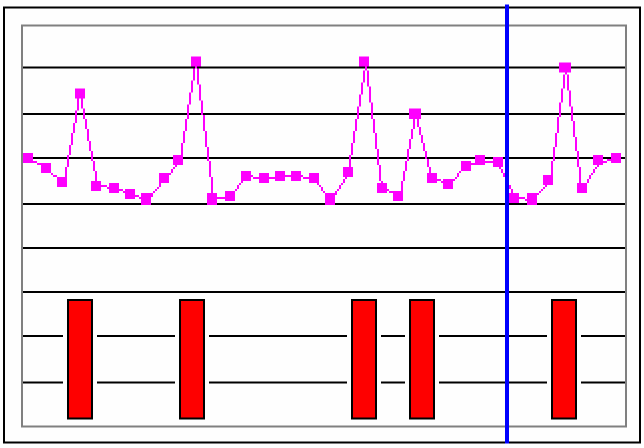

**Figure 3.** Impact of promotions/special events in a time series baseline value

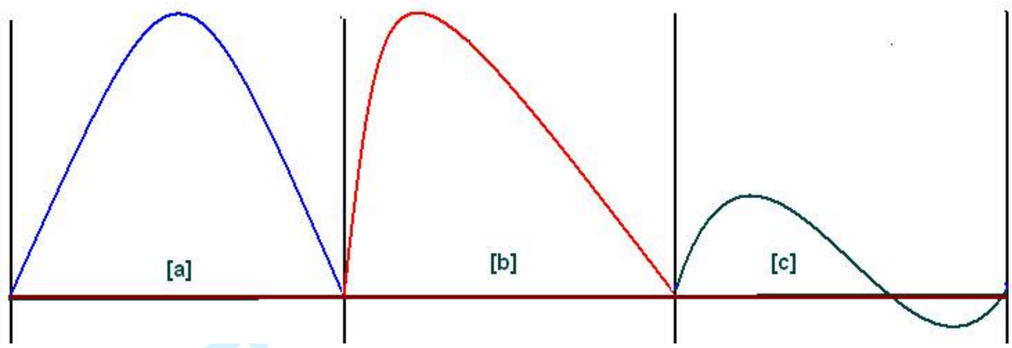

**Table 1.** Simulation Results, All data

| Methods | MAPE | MdAPE |
|---|---|---|
| MLR | 51.18% | 8.51% |
| BN | 88.05% | 21.51% |
| B3N | 113.72% | 22.36% |
| ANN | 28.03% | 7.29% |
| EXPRT1 | **25.91%** | **5.75%** |
| EXPRT2 | 26.43% | 6.21% |

**Table 2.** Simulation Results, Effect of Linearity

| Methods | Linear | | Non-linear | |
|---|---|---|---|---|
| | MAPE | MdAPE | MAPE | MdAPE |
| MLR | **9.83%** | **3.25%** | 92.53% | 25.39% |
| BN | 70.72% | 17.40% | 105.37% | 27.55% |
| B3N | 87.85% | 17.39% | 139.60% | 29.76% |
| ANN | 19.93% | 5.61% | **36.12%** | **9.51%** |
| EXPRT1 | 9.94% | 3.26% | 41.88% | 11.44% |
| EXPRT2 | 14.34% | 3.97% | 38.51% | 10.08% |

**Table 3.** Simulation Results, Effect of Noise

| Methods | High Noise | | Low Noise | |
|---|---|---|---|---|
| | MAPE | MdAPE | MAPE | MdAPE |
| MLR | 50.65% | 10.13% | 51.72% | 6.62% |
| BN | 72.95% | 21.51% | 103.15% | 21.53% |
| B3N | 85.00% | 22.37% | 142.45% | 22.36% |
| ANN | 31.75% | 8.88% | 24.30% | 5.87% |
| EXPRT1 | 27.35% | **7.01%** | 24.47% | **4.70%** |
| EXPRT2 | **29.48%** | 7.85% | **23.38%** | 4.79% |

**Table 4.** Simulation Results, Processing Time

| | Instant | 1-2 Hours |
|---|---|---|
| MLR | • | |
| BN | • | |
| B3N | • | |
| ANN | | • |
| EXPRT1 | | • |
| EXPRT2 | | • |
| *CPU : Intel P4 1.7Ghz, RAM: 512 MB* | | |

**Table 5.** Sample of Real-life Special Events Data (Source AGB Hellas S. A., adopted from Nikolopoulos *et. al.* 2007b)

| DATE | START | END | MAIN TITLE | SECONDARY TITLE | Impact | Importance | Competition | Time Zone |
|---|---|---|---|---|---|---|---|---|
| 13/5/1998 | 21:39 | 23:45 | UEFA CUP | STUTGART – CHELSEA (*FINAL)* | **2%** | 4 | 3 | 3 |
| 20/5/1998 | 21:32 | 23:50 | CHAMPIONS LEAGUE | JUVENTUS – REAL MADRID (*FINAL)* | **43%** | 4 | 2 | 3 |
| 12/7/1998 | 21:48 | 23:56 | WORLD CUP | BRAZIL - FRANCE (*FINAL)* | **38%** | 5 | 1 | 3 |

**Table 6.** Accuracy performance on Real-life Special Events Data (adopted from Nikolopoulos *et. al.* 2007b)

| Methods | MAE (Mean Absolute Error) |
|---|---|
| ANN, EXPRT1, EXPRT2 | 9.90 |
| B3N | 12.68 |
| BN | 12.89 |
| MLR | 14.54 |